UDC 711

A.E. Engelgardt

Siberian Federal University, School of Architecture and Design,
Russian Federation, Krasnoyarsk, Svobodny 79, 660041
e-mail: annengelgardt@gmail.com

BIG DATA RANKING SYSTEM AS AN EFFECTIVE METHOD OF VISUALIZING THE QUALITY OF URBAN STRUCTURAL UNITS

Abstract. *Big data is the basis for new technological changes. Constantly growing volumes of arrays greatly complicate data processing and understanding. Big data analysis extracts knowledge and meaningful information from large and complex data sets. The extraction of information displays regularities hidden in the data. Modern cities use the latest technologies to support sustainable development and a high standard of living. The indicator of a high standard of living of the urban population and, consequently, an indicator of a quality city is the quality of the urban environment. To evaluate the structural units of a city, the most common method is ranking. Ranking systems based on big data are the most effective method of visualizing the quality of structural elements of a city. Innovative ways of collecting and analyzing data are gradually replacing obsolete mechanisms of city management. Unlike statistical data, which are out of date by the time of their analysis, big data can be processed in real time that increases the quality and speed of decision making. The complexity of big data methods implementing in ranking systems is caused by problems of staff shortages, technical equipment, legal rights, security problems and openness of data. Ranking quality systems of the urban environment can be used by the city administration, designers, civil communities to assess the current state and management of the urban environment. The creation of such ranking systems is the first step towards the formation of smart open data-driven cities. The introduction of big data into cities can be divided into three levels as the influence of data on urban governance increases: applied (open data city); semi-autonomous (data-driven city); autonomous (smart city).*
Keywords: *Big data, cartography data, ranking systems, urban environment quality.*

Modern cities use the latest technologies to support sustainable development and a high standard of living for a rapidly growing population, increasing territory and increasing infrastructure. Innovative ways of collecting and analyzing, such as big data analysis, are gradually replacing obsolete mechanisms of city management.

In practice, the experience of organizations in the implementation and use of big data includes various methods and technologies for the processing, analyzing and visualizing information. They come from different areas, such as statistics, computer science, applied mathematics and economics. Examples of methods used to analyze big data are as follows: data mining- learning associative rules, classification, cluster analysis, regression analysis; crowdsourcing - categorization and enrichment of data by a broad, indefinite number of persons; mixing and data integration - a set of techniques that allow the integration of heterogeneous data from a variety of sources for the possibility of in-depth analysis; machine learning, including training with and without a teacher, as well as using models built on the basis of statistical analysis or machine learning to obtain complex forecasts based on basic models; artificial neural networks, network analysis, optimization, including genetic algorithms; spatial analysis - the use of topological, geometric and geographic information in the data; statistical analysis - A / B testing and time series analysis; visualization of analytical data - presentation of information in the form of drawings, graphs, charts and diagrams using interactive features and animations, both for results and for use as raw data for further analysis.

In addition, new analytical tools are used to analyze large data: Alterian, TweetReach (network intelligence tools for analyzing real-time reactions); NM Incite, Social Mention, SocMetrics, Traackr, Tweepi (mood analysis tools for estimating noise around a product or service, intelligent tools for identifying key people and market targeting), Attitude, Autonomy (tools for real-time testing for direct feedback on new products or ideas from users, tools for data mining, for text analytics, for estimating the size of the market).

Big Data ranking systems

Ranking systems based on big data are the most effective method of visualizing the quality of structural elements of a city.

The indicator of a high standard of living of the urban population and, consequently, an indicator of a quality city is the quality of the urban environment. To assess the structural units of the city a variety of methods, systems and sets of indicators are applied: from the most generalized indicators to detailed multi-level systems with a detailed list of components. There are such methods for assessing the quality of urban environments as comparison, classification, systematization, ranking and ranking. The most common method is the ranking method.

Ranking as a technology of comparison and ordering of research objects is now widely used in international practices for assessing the quality of cities:

- «Quality of living survey» ranking of the Mercer's international consulting company in the field of human resources;
- Annual report ranking Global Liveability Report Journal Economist Intelligence Unit, The Economist's «World's Most Liveable Cities»;
- «Most Liveable Cities Index» ranking of cities living standards by the Monocle magazine;
- «Places to Live Rankings» by the Niche's;
- «Top 100 Best Places to Live» site ranking livability;
- «Improving quality of life», «The world's top cities for people and the planet» by the Arcadis ranking company ;
- «Teleport Cities» ranking.

All these ranking systems use the open data received from various sources: from city services to social networks. To account for this amount of unstructured data, big data analysis algorithms are used. Ranking systems based on big data are the most effective method of visualizing the quality of structural elements of a city.

Data visualization using the ranking method helps analysts better understand complex data sets. And while this data visualization method often carries a conditional nature, temporary modeling can show how complex structures grow and evolve over time. (S. French, C. Barchers W., Zhang, 2015). Urban planning is mainly focused on the long term - what is happening in cities is measured in months or years, not hours and minutes, the exception is only the planning of transport (daily flows). However, to understand how cities function during the day is of great value. (Michael Batty, 2013)

Unlike statistical data that can be outdated by the time of their analysis, big data can be processed in real time that increases the quality and speed of decision-making. Dynamic quantitative assessment helps to see the overall picture of cities and the trends of their development. The combination of a spatial component and a dynamic sensory base seems to be the only way to meet the modern requirements of urban management. (Juan R. Selva-Royo, Nuno Mardones, Alberto Cendoya, 2017). Thus, the ranking systems of urban environment quality can be used by the city administration, the designers not only to assess the current state, but also to manage the urban environment. However, there are certain difficulties in implementing methods for analyzing large data for their further use in ranking systems.

Problems and difficulties of Big Data ranking systems implementation

Big data as an area associated with information technology has its own set of difficulties and problems.

In developed countries, information is the most significant resource. Under these conditions, the ability to collect the right data and information and transform it into effective and useful knowledge is becoming an increasingly important requirement. While data processing, ownership and storage have become almost free, and networks and cloud storage provide users with global access and ubiquitous services, many new problems have simultaneously occurred (Janusz Wielki, 2013). There are five main problems:

1. Staff shortage: there are no skilled experienced people (Data scientists) able to work with Big Data.

2. Technical equipment: big data analysis is superior to traditional methods in terms of processing speed, but it is still a long process when it comes to huge data sets.

3. Openness of data: the dynamic nature of large data methods requires the creation of open access platforms, where data will be received as transparent and impartial as possible. (Ma, X., Wang, Y., 2014) To date, many data of city management companies are closed to free access that makes getting information more difficult.

4. Legal Rights: the potential of using social networking resources is huge. On the planet there are about 7 billion people who access 1.2 billion personal computers (as of 2013) and their number is constantly growing, as well as the amount of data they create. (Michael Batty, 2013) However, there are problems such as copyright, database rights, confidentiality, trademarks, contract law, competition law. There is another important challenge that is related to legal aspects - this refers to transparency in the practice of data collection.

5. Security problem: at the same time, existing big data problems and threats are still developing, for example, the problem of ensuring the security of collected data and information. These issues mainly concern the protection of confidential data and data that should be stored by private organizations (for example, various types of consumer data). Large data can reveal the most personal aspects of our behavior, who we visit, what we buy and so on. As a result, problems related to the security of IT organizations, IT infrastructures and protection from various attacks are becoming even more important than before. (Tene, O., & Polonetsky, J., 2012) New organization systems will generate a high degree of dependence on data, it will be necessary to monitor the reliable functioning of the IT infrastructure continuously, since in case of a failure, important city management centers will cease to function.

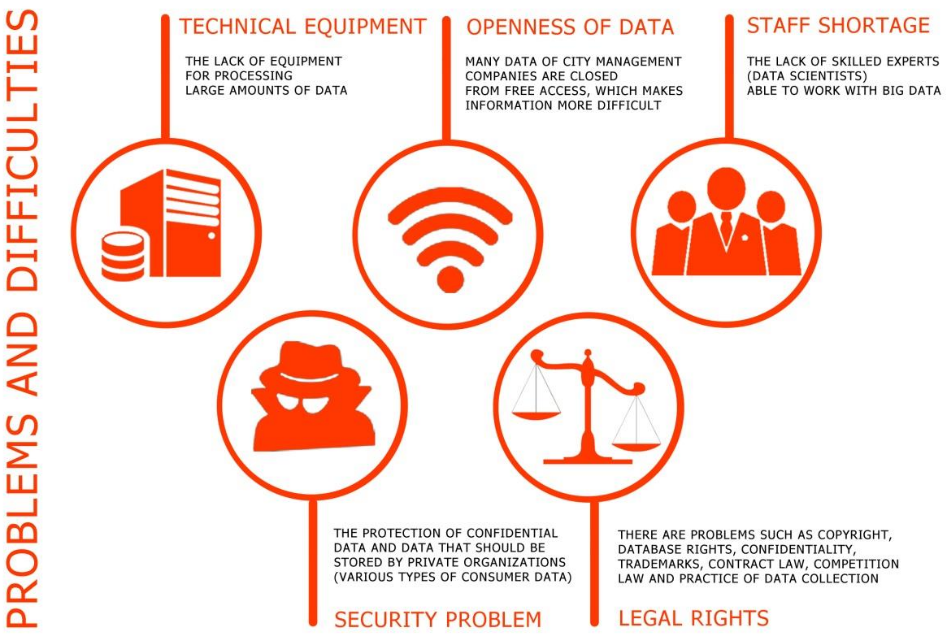

Figure 1. Problems and difficulties of Big Data implementation

Recommendations for Big Data Ranking systems implementing

During the analysis of ranking systems, the following recommendations for their implementation have been proposed:

1. Local level: ranking systems based on big data are mainly presented at the global level (a comparison of countries and cities among themselves), but these systems are only a reflection of the situation as a whole and can not point to specific problems within urban areas. Therefore, the introduction of these rankings at the local level (between administrative areas and structural units within the city) may be more efficient for urban planning

2. Real time mode: the frequency of updating rankings basically makes an annual interval. In an ever-changing world, the information obtained as a result will not be sufficiently reliable. Systems should display the situation in real time.

3. Spatial data modeling assumes methods for visualizing large data.

4. Inclusion of the population in the analysis: spatial modeling using big data is already an established method of research, but public participation in it is a new area that involves good prospects. (J. Hao, 2015)

On the basis of the Programming Processing environment, the supposed image of "Interactive ranking map for the quality of structural units of the Krasnoyarsk city" program has been developed - this is a project proposal of a simulation model of a ranking system based on big data.

The given project assumes the creation of a ranking at the local level. The peculiarity of the model is that the ranking can be displayed both at the administrative level and the quality of the structural units of the city outside the administrative division. The availability and openness of the ranking for citizens is expected to provide the following:

1. Creation of a website of a city ranking, updated in the real time;

2. Venues for interactive participation in urban planning (by registering on the site and using planning tools for the population;

3. Visualization of the current condition of structural and administrative units

Figure 2. Interactive ranking map for the quality of structural units of the Krasnoyarsk city

Urban Big Data

The introduction of ranking systems based on big data in our cities can improve the efficiency of urban policies. The creation of these ranking systems is the first step towards the formation of smart open data-driven cities.

If you combine structured data (geospatial and telecommunication data, logistics data, data from measuring devices) obtained from urban planning with unstructured data streams from social networks and websites (all posted videos, images and text data that contain timestamps and locations, individual photos, video clips and intelligent systems), channels from surveillance cameras, unmanned aerial vehicles, then we will get big urban data. (S. French, C. Barchers W., Zhang Moving, 2015)

Big urban data allow building models and testing town planning theories, and they are a modern and effective method of urban planning (Fig. 1). Examples of such planning are the determination of the density of flows on specific sections of roads, determination of the dynamics of urban activities (real-time cities) and migration flows of the population during the day, determining the actual girth of the city (Juan R. Selva-Royo, Nuno Mardones, Alberto Cendoya, 2017)

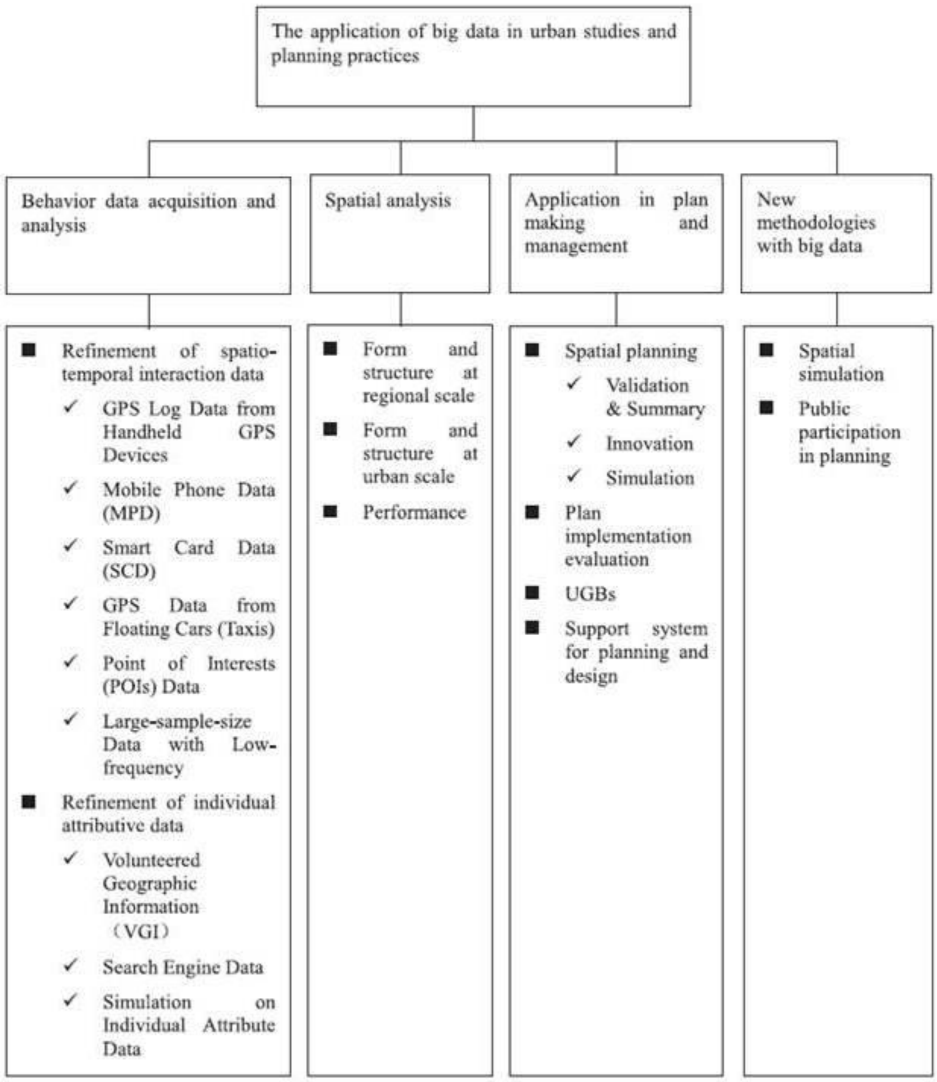

Figure 3. J. Hao et al. / Journal of Urban Management 4 (2015) 92–124

The use of big data analysis in urban planning introduces structural changes to cities, since this method involves the optimization of ecological, economic and social models of urban systems. When studying the world experience of introducing big data into cities, it is possible to identify three levels of interaction as the influence of data on urban governance increases:

Level 1 – an pplied level (open data city);

Level 2 – a semi-autonomous level (data-driven city);

Level 3 – an autonomous level (smart city).

At the application level, all urban infrastructure systems receive continuous reports on their condition, buildings transmit energy cost data, water and air quality sensors provide real-time information about dozens of environmental parameters; cell phone location data provide a detailed overview of the activity of millions of urban residents. All data is freely available and open to the public in real time. Most of the large multimillion cities open data and can be assigned to this category.

The semi-autonomous level implies the integration of data flows into single integrated systems and the creation of data management centers. At this level, the obtained data is the determining factor of urban governance. A centralized management system can learn and improve over time with the help of machine learning methods. When machine learning methods are applied to large databases, such as big data, they are called intelligent data. Data collection attempts to identify and build a simple model with high prediction accuracy, based on a large amount of data. Then the model is used to predict the future value. The city of São Paulo in Brazil can be taken as a sample of a city managed by data, in 2014 the Department of Public Security of the state announced the launch of a system developed in partnership with Microsoft called Detecta. The project was inspired by a similar initiative that has been working in New York since 2012 and is to provide data to the police headquarters that come from thousands of phone calls, video cameras located throughout the city and thousands of policemen patrolling the streets into their contact centers. There are similar centers in the city of Rio de Janeiro.

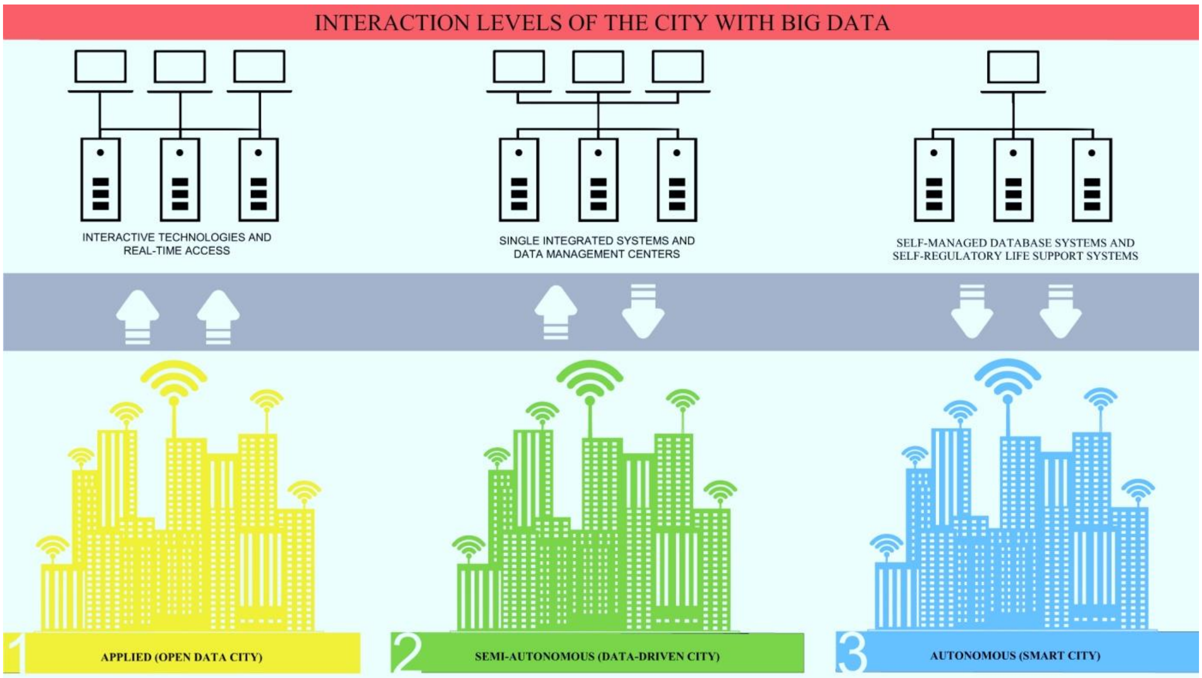

Figure 4. Interaction levels of the city with big data

Smart cities are an example of the autonomous level that is characterized by the emergence of self-managed database systems and self-regulatory life support systems. An intelligent city, an information city, a virtual city can be used as synonyms for the term of an intelligent city. (Michael Batty, 2013) The key principle here is the data reuse that means the data of IT resources are available for use by other systems. Those cities include: Chicago, US; Yinchuan, China; Fujisawa, Japan; Milton Keynes, UK; Singapore; Masdar, United Arab Emirates.

Given the ever-increasing volumes of information arrays, as well as the complexity of the data, the importance of implementing methods of big data analysis in urban management will only increase in the future years.

Big data make it possible to understand the nuances of urban systems and to find a correlation and causality that cannot be found in random sampling through small surveys. Now you can get a detailed idea of how people, businesses and institutions use the urban space and more quickly identify complex patterns of the behavior that characterize the urban area. However, it is important to understand that big data cannot replace the theoretical foundation needed to plan an uncertain future, since big data analysis is a way of short-term management.

Conclusion

Big data analysis is a new mechanism of city management. Big data ranking systems are the most effective method of visualizing the quality of structural elements of a city. Unlike statistical data that can be outdated by the time of their analysis, big data can be processed in real time that increases the quality and speed of decision-making. Dynamic quantitative assessment helps to see the overall picture of cities and trends in their development.

Technologies and methods for big data are within reach, but the complexity of their implementation is caused by problems of staff shortages, technical equipment, legal rights, security problems and openness of data.

The introduction of ranking systems based on big data largely depends on the determination of the city authorities to change the approach to urban planning. The use of large data analysis in urban planning introduces structural changes to cities, since this method involves the optimization of urban systems. Big urban data allow building models and testing urban theories.

The introduction of large data into cities can be divided into three levels as the influence of data on urban governance increases: an applied level (open data city); a semi-autonomous level (data-driven city); an autonomous level (smart city).

References

1. Aminreza Iranmanesh, Resmiye A. Atun. (2017) *Exploring patterns of socio-spatial interaction in the public spaces of city through Big Data.*
2. Awais Latif, P., Kammeier D., H., & Liu, H. (2003). *Planning-support system as an innovative blend of computer tools: An approach for guiding decisions on industrial locations in Punjab Province*, Pakistan. Urban Planning Overseas, 18(5), 15–20.
3. Bibri, S. E., Krogstie, J. (2017) *The core enabling technologies of big data analytics and context‑aware computing for smart sustainable cities: a review and synthesis*, S. Bibri and Krogstie J., Journal of Big Data.
4. Bibri SE . (2017) *Smart sustainable cities of the future: the untapped potential of big data analytics and context–aware computing for advancing sustainability*. Urban Book Series. Springer; (in press).
5. Boyd, D., & Crawford, K. (2012). *Critical questions for big data: Provacations for a cultural, technological, and scholarly phenomenon. Information, communication & society*, 15(5). 662-679.
6. Chai, Y., Shen, Y.,& Chen, Z.(2014).*Towards smarter cities: Human-oriented urban planning and management based on space-time behavior research*. Urban Planning International, 29(6), 31.37þ50.
7. Chang, X., Yue, Y., Li, Q., Chen, B., Shaw, S., & Tu, W. (2014). *Extracting the geographic backbone of location –based social network*. Geomatics and Information Science of Wuhan University, 39(6).
8. Chen, J., Shaw, S., Yu, H., Lu, F., Chai, Y., & Jia, Q. (2011). *Exploratory data analysis of activity diary data: A space-time GIS approach.* Journal of Transport Geography, 19(3), 394–404.
9. Graham, M., & Shelton, T. (2013). *Geography and the future of big data, big data and the future of geography*. Dialogues in Human Geography, 3(3), 255-261.
10. Crampton, J.W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M., Wil-son, M. W., & Zook, M. (2013). *Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb*. Cartography and geographic Information science, 40(2). 130-139.
11. Dang, A., Yuan, M., Shen, Z., & Wang, P. (2015). *Reflections on rational planning and urban & rural governance based on the conception of smart city and big data*. Construction Science and Technology, 5, 64–66.
12. DeRen L, JianJun C, Yuan Y. (2015) *Big data in smart cities*. Sci China Inf Sci.;58:1–12.
13. French S., Barchers C., Zhang W. (2015) *Moving beyond Operations: Leveraging Big Data for Urban Planning Decisions. CUPUM.*
14. Gurban I.A. (2015) *Ranking of territories as a tool for measuring regional well-being* / UDC 338.001.36 (in Russian)

15. Inostroza, L. et al. (2013) Urban sprawl and fragmentation in Latin America: A dynamic quantification and characterization of spatial patterns. Journal of Environmental Management 115, 87–97.
16. Janusz Wielki. (2013) *Implementation of the Big Data concept in organizations –possibilities, impediments and challenges /* Proceedings of the 2013 Federated Conference on Computer Science and Information Systems pp. 985–989.
17. Jin, X., Zhang, Z., Wang, B., & Zhu, X. (2015).*Reflections of informatization of planning in big data era*. Planners, 31(3), 135–139.
18. Jinwei Hao, Jin Zhu, Rui Zhong. (2015) *The rise of big data on urban studies and planning practices in China: Review and open research issues*. Journal of Urban Management 492–124.
19. Selva-Royo, J. R., Mardones, N., Cendoya, A. (2017) *Cartographying the real metropolis: A proposal for a data-based planning beyond the administrative boundaries.*
20. Katal A, Wazid M, Goudar R. (2013) *Big Data: issues, challenges, tools and good practices*. In: Proceedings of 6th International Conference on Contemporary Computing (IC3), Noida, August 8–10, IEEE, US. pp. 404–409.
21. Kumar A, Prakash A. (2014) *The role of big data and analytics in smart cities*. Int J Sci Res (IJSR);6(14):12–23.
22. Li, X., Li, D., Liu, X., & He, J. (2009). *Geographical simulation and optimization system (GeoSOS) and its cutting-edge researches.* Advances in Earth Science, 24(8), 899–907.
23. Li, C. (2014). *Study on the regional space of flows based on location data from social network: A case study of city group of Yangtze River Delta*. Shanghai Urban Planning Review, 5, 44–50.
24. Li, M., & Wang, P.(2014). *New technology in data-driven urban planning: From GIS to Big data*. Urban Planning International, 29(6), 54–65.
25. Long, Y., Shen, Z., Mao, Q., & Hu, Z. (2011). *Urban growth control planning support system: Approach, implementation and application*. City Planning Review, 35(3), 62–71.
26. Long, Y., Mao, M., Mao, Q., Shen, Z., & Zhang, Y.(2014).*Fine-scale urban modeling and its opportunities in the "big data" ear: Methods, data and empirical studies*. Human Geography, 29(3), 7–13.
27. Long, Y., Wu, K., Wang, J., & Liu, X. (2014). *Big models: A novel paradigm for urban and regional studies*. Urban Planning Forum(6), 52–60.
28. Ma, X., & Wang, Y. (2014). *Development of a data-driven platform for transit performance measures using smart card and GPS data*. Journal of Transportation Engineering, 140(12), 04014063.
29. Mayer-Schoenberger V., Kukier K.. (2015) *Big data. A revolution that will change how we live, work and think*. (in Russian)
30. Batty, M.. (2013) *Big data, smart cities and city planning*
31. Ming-Chun Lee. (2017) *Case study on emerging trends in geospatial technologies for study of urban form*
32. Nathalie, M., Symeon, P., Antonio P, Kishor T. . (2012) *Combining cloud and sensors in a smart city environme*nt. EURASIP J Wirel Commun Netw.; 247:1–10.
33. Qin, X., & Zhen, F.(2014). *The spatial planning methods of smart city on big data era*. Modern Urban Research, 10, 18–24.
34. Pérez-delHoyo, R., Lees, M. C. (2017) *Redefining the Smart City concept: the importance of humanizing 'Intelligent' Cities.*
35. Shin D. (2009) *Ubiquitous city: urban technologies, urban infrastructure and urban informatics*. J Inf Sci.,;35(5):515–26.
36. Tene, O., & Polonetsky, J. (2012*) Big data for all: Privacy and user control in the age of analytics*. Nw. J. Tech & Intell. Prop., 11, xxvii.
37. Townsend A. (2013) *Smart cities—big data, civic hackers and the quest for a new utopia.* New York: Norton & Company.
38. Zhen, F., & Qin, X.(2014). *The application of big data in smart city research and planning.* Urban Planning International, 29(6), 44–50.

Illustrations